

## Analog VLSI for Auditory and Vision Signal Processing

Carver A. Mead

California Institute of Technology

The visual system of a single human being does more image processing than the entire world's supply of supercomputers. Digital computers are extremely effective at producing precise answers to well-defined questions, but are next to useless when the input is fuzzy and poorly conditioned, the computation is not precisely defined, and the output need be only approximately "correct". Part of the reason for this large disparity in capabilities is that a large proportion of neural computation is done in an *analog* rather than a digital manner. Rather than view an electrical signal as a single boolean variable, the entire range of signal values is used to represent information. Instead of preconceived **and** and **or** functions, the nervous system uses elementary computational primitives that are a direct consequence of fundamental laws of physics. This clever choice of primitives gives neural tissue many orders of magnitude advantage in energy dissipated per operation.

We are naturally led to ask if we can apply this same trick to systems built of silicon. After all, the underlying physics of the two media is, when viewed with the proper warped perspective, very similar. Signals are represented as differences in electrical potential, and are conveyed on "wires" formed by surrounding a conducting path with an excellent electrical insulator. Active devices cause electrical current to flow in a second "output" conductor due to the potential in a first "input" conductor. The "output" of an active device has more energy than was present in the "input" to that device; and hence the systems possess "gain"—the essential ingredient for unbounded information processing—which is accompanied by an unavoidable dissipation of energy. A "power supply" maintains a near-constant average difference in electrochemical potential across the active devices. The active devices are formed of extremely thin *energy barriers* which prevent the flow of current between two electrical nodes. The passage of current is mediated by the potential on a third "control" electrical node. That current varies exponentially with the potential on the control node.

In the subthreshold regime, a single MOS transistor produces a drain current that is exponential in the input voltage. This exponential function is a direct result of the Boltzmann distribution. Addition and subtraction of currents follows directly from the conservation of charge (Kirchhoff's law). Temporal integration is the elementary property of a capacitor. Out of these three primitives, we can construct a computational paradigm of far greater relevance to certain classes of problems than its digital counterpart.

The rapid development of CMOS integrated-circuit technology has given us a medium in which it is presently possible to fabricate tens of millions of devices interconnected on a single silicon wafer. This number will increase by two orders of magnitude before fundamental limitations are encountered. This technology is popularly conceived as a digital technology, and has been highly evolved for the production of microprocessors, memories, and other digital products. It might therefore be supposed that the most highly evolved fabrication process would not be suitable for the functions required in neural processing.

The noise properties of a typical device, or the precision with which any two devices can be matched, for example, are far worse than technologies historically used for implementing analog functions. We observe, however, that the precision, reliability, and noise properties available in neural wetware fall short of that used in even the most rudimentary electronic systems. This lack of precision and reproducibility at the component level is more than offset by the redundancy introduced at the system level. Whether this property is the primary reason for the large connectivity in neural systems, or whether it is a byproduct of an organizing principle dictated by other system needs, is not a question open to us at the present time. We do know, however, that robustness under failure or imprecision of individual components is one important emergent property of neural systems. If we base our designs on the same organizing principles, we should not be concerned that individual devices will cause system malfunction. To the contrary, we can expect that systems with extraordinary reliability and robustness will result; so much so that useful integration at the scale of a complete wafer is feasible. These themes are the subject of the new book *Analog VLSI and Neural Systems*, Addison-Wesley, 1988.

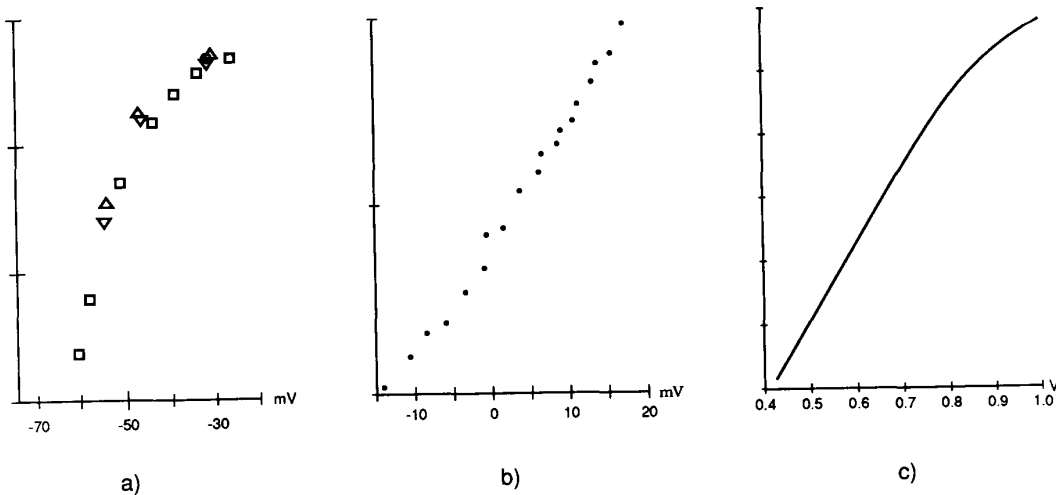


Fig.1. Exponential current-voltage relations found in nature. In all cases, the vertical axis is a logarithmic measure of current; Each division represents a factor of 10. Part a) is the characteristic of the sodium conductance in the giant axon of the squid. Part b) is the transfer characteristic of a chemical synapse. Part c) is the saturation current of a typical MOS transistor.